

Antenna Selection for MIMO Non-orthogonal Multiple Access Systems

Yuehua Yu, He Chen, Yonghui Li, Zhiguo Ding, and Branka Vucetic

Abstract

This paper considers the joint antenna selection (AS) problem for a classical two-user MIMO non-orthogonal multiple access (NOMA) system, where both the base station (BS) and users (UEs) are equipped with multiple antennas. Specifically, several computationally-efficient AS algorithms are developed for two commonly-used NOMA scenarios: fixed power allocation NOMA (F-NOMA) and cognitive radio-inspired NOMA (CR-NOMA). For the F-NOMA system, two novel AS schemes, namely max-max-max AS (A^3 -AS) and max-min-max AS (AIA-AS), are proposed to maximize the system sum-rate, without and with the consideration of user fairness, respectively. In the CR-NOMA network, a novel AS algorithm, termed maximum-channel-gain-based AS (MCG-AS), is proposed to maximize the achievable rate of the secondary user, under the condition that the primary user's quality of service requirement is satisfied. The asymptotic closed-form expressions of the average sum-rate for A^3 -AS and AIA-AS and that of the average rate of the secondary user for MCG-AS are derived, respectively. Numerical results demonstrate that the AIA-AS provides better user-fairness, while the A^3 -AS achieves a near-optimal sum-rate in F-NOMA systems. For the CR-NOMA scenario, MCG-AS achieves a near-optimal performance in a wide SNR regime. Furthermore, all the proposed AS algorithms yield a significant computational complexity reduction, compared to exhaustive search-based counterparts.

Index Terms

Multiple-input multiple-output (MIMO), non-orthogonal multiple access (NOMA), joint antenna selection (AS), cognitive radio.

The material in this paper has been submitted in part at the IEEE International Conference on Communications, Paris, France, 2017 [1].

Y. Yu, H. Chen, Y. Li, and B. Vucetic are with the School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia (email: {yuehua.yu, he.chen, yonghui.li, branka.vucetic}@sydney.edu.au).

Z. Ding is with the School of Computing and Communications, Lancaster University, Lancaster LA1 4YW, U.K. (email: z.ding@lancaster.ac.uk).

I. INTRODUCTION

The non-orthogonal multiple access (NOMA) technique has been emerged as a promising solution to significantly improve the spectral efficiency of the next-generation wireless networks [2]–[4]. By superimposing the information of multiple users (UEs) in the power domain, multiple UEs can be served within the same time, frequency and code domain. Different from the conventional water-filling power allocation strategy, the NOMA technique generally allocates more transmit power to the UEs with the poor channel conditions (i.e., *weak* UEs). In this case, these UEs can decode their higher-power-level signals directly by treating others' signals as noise. In contrast, those UEs with the better channel conditions (i.e., *strong* UEs) adopt the successive interference cancellation (SIC) technique for signal detection. Specifically, the *weak* UEs' messages are first decoded and subtracted before their own signals with lower-power-levels can be recovered. It has been demonstrated that both the throughput performance and user fairness can be significantly improved in NOMA systems compared to conventional orthogonal multiple access (OMA) systems [2], [5].

Initial efforts on the design and analysis of the NOMA technique focused on single-antenna systems, see e.g., [6]–[10]. Recently, multiple antennas have been used in NOMA systems (MIMO-NOMA) to exploit the spatial degrees of freedom [11]–[15]. Although the capacity performance can potentially scale up with the number of antennas, this superior performance comes at the price of expensive RF chains at all terminals. To avoid the high hardware costs while preserving the diversity and throughput benefits from MIMO, the antenna selection (AS) technique has been recognized as an effective solution [16]–[18]. There are only a few papers that considered the AS problem for MIMO-NOMA systems in the open literature. Specifically, a transmit AS (TAS) algorithm was proposed in [19], and a joint TAS and user scheduling algorithm was considered in [20]. However, both [19] and [20] only focused on the TAS design at the base station (BS) side as each UE was assumed to be equipped with a single antenna. Moreover, there have been no analytical characterization of the system performance for algorithms in [19] and [20].

To our best knowledge, joint AS at both the BS and UEs for MIMO-NOMA systems is still an open problem. Though there have been some published results on joint AS schemes in conventional MIMO-OMA systems, they cannot be extended to MIMO-NOMA systems directly. This is mainly because there is severe inter-user interference in MIMO-NOMA systems while the signals are transmitted in an interference-free manner in MIMO-OMA systems. A global optimal solution of this problem requires an exhaustive search (ES) over all possible antenna combinations, and its complexity would become unacceptable when the numbers of antennas at both the BS and UEs become large. Motivated by this, in this paper we propose several low-complexity joint AS algorithms for two commonly used NOMA scenarios:

fixed power allocation NOMA (F-NOMA) and cognitive radio-inspired NOMA (CR-NOMA) [21]–[23]. Specifically, the classical two-user NOMA model is adopted, which is how NOMA is implemented in LTE-Advanced. For the F-NOMA scenario, two novel joint AS algorithms, namely max-max-max AS (A^3 -AS) and max-min-max AS (AIA-AS), are proposed to maximize the system sum-rate without and with the consideration of user fairness, respectively. Specifically, A^3 -AS aims to maximize the rate performance of the instantaneous *strong* UE, while AIA-AS tries to improve the rate performance of the instantaneous *weak* UE. In the CR-NOMA system, two UEs have different priorities: one is treated as the primary UE and the other is treated as the secondary UE. Moreover, the secondary UE is opportunistically served only when the primary UE's quality of service (QoS) is guaranteed. That is, the BS transmit power allocated to the secondary UE is constrained by the primary UE's instantaneous signal-to-interference-plus-noise ratio (SINR). For this case, a novel joint AS algorithm, referred to as maximum-channel-gain-based AS (MCG-AS), is proposed to maximize the achievable rate of the secondary UE in CR-NOMA systems. Furthermore, two simplified versions of MCG-AS, namely primary-user-based AS (PU-AS) and secondary-user-based AS (SU-AS), are designed to further reduce the computational complexity of MCG-AS. The main contributions of the paper are summarized as follows:

- Two joint AS algorithms (i.e., A^3 -AS and AIA-AS) are proposed to maximize the system sum-rate of F-NOMA systems without and with the consideration of user fairness, respectively;
- A joint AS scheme (i.e., MCG-AS) and its two simplified versions (i.e., PU-AS and SU-AS) are proposed to maximize the achievable rate of the secondary UE in CR-NOMA systems subject to the QoS requirements of the primary UE;
- The computational complexities for all the proposed joint AS algorithms are considered, and are shown to be much lower than that of the optimal ES-based scheme;
- Asymptotic closed-form expressions for the average sum-rates of both A^3 -AS and AIA-AS in F-NOMA systems are derived for high SNR scenarios;
- Asymptotic closed-form expressions for the average rate of the secondary UE of MCG-AS, PU-AS and SU-AS in CR-NOMA systems are derived for high SNR scenarios;
- Extensive simulation results illustrate that both A^3 -AS and AIA-AS in F-NOMA systems yield significant performance gains over the orthogonal multiple access ES scheme (OMA-ES) and the random selection scheme. In particular, A^3 -AS can achieve the near-optimal system sum-rate while AIA-AS can provide better user fairness. In CR-NOMA, the MCG-AS scheme can achieve the near-optimal rate performance in a wide SNR regime. In contrast, the rate performance of both PU-AS and SU-AS is significantly affected by the distances between the BS and UEs. Specifically, the PU-AS

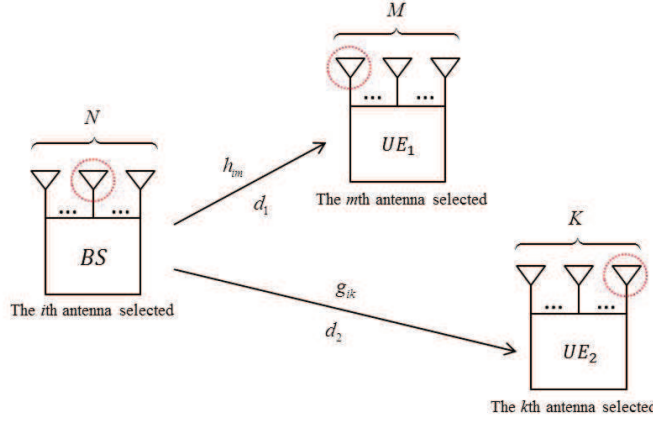


Fig. 1: Diagram of a two-user MIMO-NOMA down-link system.

(SU-AS) scheme can approach the optimal performance when the primary (secondary) UE is closer to the BS.

The rest of the paper is organized as follows. In Section II, the considered two-user MIMO-NOMA down-link model is introduced, and the joint AS optimization problems for both F-NOMA and CR-NOMA systems are formulated, respectively. In Section III, two novel joint AS schemes for F-NOMA systems are proposed and their achievable sum-rate performances are analyzed. In Section IV, a computationally-efficient joint AS algorithm and its two simplified versions are developed and analyzed for CR-NOMA systems. Finally, simulation results and comparisons between the proposed algorithms and benchmark schemes are given in Section V, and conclusions are drawn in Section VI.

Notations: $\mathcal{C}^{m \times n}$ represents the set of all $m \times n$ matrices. All bold uppercase letters represent matrices, and all Calligraphic letters represent sets. $\Pr(\cdot)$ denotes the probability of an event, and $|\cdot|$ and $\mathbb{E}[\cdot]$ denote the absolute value and the expectation operation, respectively. $F_X(x)$ and $f_X(x)$ represent the cumulative density function (CDF) and the probability density function (PDF) of a random variable X , respectively. \mathcal{O} is usually used in the efficiency analysis of algorithms and $q(x) = \mathcal{O}(p(x))$ when $\lim_{x \rightarrow \infty} \left| \frac{q(x)}{p(x)} \right| = c, 0 < c < \infty$.

II. SYSTEM MODEL

Consider a two-user¹ MIMO-NOMA down-link scenario as [12]. As shown in Fig. 1, the BS is equipped with N antennas, while user one (UE₁) and user two (UE₂) are equipped with M and K antennas, respectively. The channel matrix from the BS to UE₁ is denoted by $\mathbf{H} \in \mathcal{C}^{N \times M}$ and that from the BS to

¹The two-user setting is how NOMA is implemented in LTE-Advanced.

UE₂ is denoted by $\mathbf{G} \in \mathcal{C}^{N \times K}$. We assume that the channels between the BS and UEs undergo spatially independent flat Rayleigh fading. The entries of \mathbf{H} (\mathbf{G}), e.g., \tilde{h}_{im} (\tilde{g}_{ik}), can be modeled as independent and identically distributed (i.i.d.) complex Gaussian random variables, where \tilde{h}_{im} (\tilde{g}_{ik}) represents the channel coefficient between the i th antenna of the BS and the m th (k th) antenna of UE₁ (UE₂). For notation simplicity, we define $h_{im} = |\tilde{h}_{im}|^2$ and $g_{ik} = |\tilde{g}_{ik}|^2$.

We consider that all the nodes are equipped with one RF chain due to the cost consideration. In each resource block, the BS selects one (e.g., i th) out of N available antennas to transmit information, while the UEs select one (e.g., m th and k th) out of M and K available antennas to receive messages, respectively. The global information of the channel amplitudes is assumed to be perfectly known at the BS through the control signalling.

Let δ denote the channel order indicator, defined as

$$\delta = \begin{cases} 1, & \text{if } h_{im} \geq g_{ik}, \\ 0, & \text{if } h_{im} < g_{ik}. \end{cases} \quad (1)$$

According to the NOMA principle, the BS broadcasts the signals superimposed in the power domain as

$$x = \delta(\sqrt{b}s_1 + \sqrt{a}s_2) + (1 - \delta)(\sqrt{a}s_1 + \sqrt{b}s_2), \quad (2)$$

where s_i denotes the signal to UE _{i} with $\mathbb{E}[|s_i|^2] = 1$, and a and b are the power allocation coefficients satisfying $a + b = 1$. Without loss of generality, we assume that $a > b$ is set to guarantee that more power is allocated to the instantaneous *weak* UE.

The received signals at UEs are given by

$$y_1 = \sqrt{P_s} \tilde{h}_{im} x + n_1, \quad (3)$$

$$y_2 = \sqrt{P_s} \tilde{g}_{ik} x + n_2, \quad (4)$$

where P_s is the transmit power at the BS, and n_i is the complex additive white Gaussian noise (AWGN) with variance σ_i^2 . For notational simplicity, we hereafter assume $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

When $\delta = 1$, UE₂ is the *weak* user and UE₁ is the *strong* user, hence the power level of s_2 is larger than that of s_1 . In this case, UE₂ decodes s_2 directly by treating s_1 as noise. In contrast, UE₁ first decodes s_2 and subtracts it by performing SIC, then decodes its own s_1 without interference. For the case $\delta = 0$, the decoding order is reversed. By using the fact that the channels are ordered, it can be easily verified that SIC can be implemented successfully, and the following two rates are achievable to UE₁ and UE₂,

respectively :

$$R_1 = \delta \log_2 (1 + \rho b h_{im}) + (1 - \delta) \log_2 \left(1 + \frac{a h_{im}}{b h_{im} + \frac{1}{\rho}} \right), \quad (5)$$

$$R_2 = \delta \log_2 \left(1 + \frac{a g_{ik}}{b g_{ik} + \frac{1}{\rho}} \right) + (1 - \delta) \log_2 (1 + \rho b g_{ik}), \quad (6)$$

where $\rho = P_s/\sigma^2$ is the transmit SNR.

In the following subsections we formulate the joint AS optimization problems for both F-NOMA and CR-NOMA systems, respectively.

A. Joint AS Optimization Problem for F-NOMA Systems

In F-NOMA systems, fixed power allocation coefficients (i.e., a and b) are assumed and both UEs have the same priority to access the network. In this case, we can formulate the following optimization problem to maximize the achievable system sum-rate

$$\mathbf{P1} : \{i^*, m^*, k^*\} = \arg \max_{i \in \mathcal{N}, m \in \mathcal{M}, k \in \mathcal{K}} R_{\text{sum}}(h_{im}, g_{ik}), \quad (7)$$

where $\mathcal{N} = \{1, 2, \dots, N\}$, $\mathcal{M} = \{1, 2, \dots, M\}$ and $\mathcal{K} = \{1, 2, \dots, K\}$. Besides, the achievable sum-rate of the system is given by

$$R_{\text{sum}}(h_{im}, g_{ik}) = R_1 + R_2 = \log_2 (1 + \rho b \gamma_i^s) + \log_2 \left(1 + \frac{a \gamma_i^w}{b \gamma_i^w + 1/\rho} \right), \quad (8)$$

where $\gamma_i^s = \max(h_{im}, g_{ik})$ denotes the channel gain of the instantaneous *strong* UE and $\gamma_i^w = \min(h_{im}, g_{ik})$ denotes the channel gain of the instantaneous *weak* UE.

B. Joint AS Optimization for CR-NOMA Systems

In the considered two-user CR-NOMA down-link scenario, without loss of generality, we treat UE₂ as the primary user and UE₁ as the secondary user. Specifically, UE₁ is opportunistically served when the QoS of UE₂ is achieved. Hereafter we use R_1^c and R_2^c to respectively denote the achievable rate of UE₁ and UE₂, where the superscript $(\cdot)^c$ is used to distinguish the parameters in CR-NOMA systems from those in F-NOMA systems. Mathematically, the achievable rate of the secondary user, R_2^c , should satisfy $R_2^c \geq R_{\text{th}}$, where R_{th} is the QoS threshold of UE₂. By noting that the power allocation coefficient $0 \leq b^c \leq 1$, we then can express the valid range of b^c as below:

$$b^c = \begin{cases} \min\left(\frac{\varepsilon}{\rho g_{ik}}, 1\right), & \text{if } \delta = 0, \\ \max\left(\frac{\rho g_{ik} - \varepsilon}{\rho g_{ik}(\varepsilon + 1)}, 0\right), & \text{if } \delta = 1, \end{cases} \quad (9)$$

where $\varepsilon = 2^{R_{th}} - 1$, and we then have $a^c = 1 - b^c$.

By substituting (9) into (5), we can observe that the achievable rate of the secondary user, R_1^c , is zero when $\delta = 0$, $b^c = 1$ or $\delta = 1$, $b^c = 0$. In other words, when the QoS of UE₂ cannot be satisfied, no power would be allocated to UE₁ and hence $R_1^c = 0$. For the considered high SNR scenarios (i.e., when $\rho \rightarrow \infty$), the expression of b^c given in (9) can be simplified as

$$b^c = \begin{cases} \frac{\varepsilon}{\rho g_{ik}}, & \text{if } \delta = 0, \\ \frac{\rho g_{ik} - \varepsilon}{\rho g_{ik}(\varepsilon + 1)}, & \text{if } \delta = 1. \end{cases} \quad (10)$$

By substituting (10) into (5), the achievable rate of the secondary user (i.e., R_1^c) can be represented as

$$\begin{aligned} R_1^c(h_{im}, g_{ik}) &= \delta \log_2 \left(1 - \frac{\varepsilon h_{im}}{(\varepsilon + 1)g_{ik}} + \frac{\rho h_{im}}{\varepsilon + 1} \right) + (1 - \delta) \log_2 \left(\frac{g_{ik}}{\varepsilon h_{im} + g_{ik}} + \frac{\rho h_{im} g_{ik}}{\varepsilon h_{im} + g_{ik}} \right) \\ &\stackrel{\rho \rightarrow \infty}{\approx} \delta \log_2 \left(\frac{\rho h_{im}}{\varepsilon + 1} \right) + (1 - \delta) \log_2 \left(\frac{\rho h_{im} g_{ik}}{\varepsilon h_{im} + g_{ik}} \right). \end{aligned} \quad (11)$$

Now we can formulate the following joint AS problem for CR-NOMA systems:

$$\begin{aligned} \mathbf{P2} : \{i^*, m^*, k^*\} &= \arg \max_{i \in \mathcal{N}, m \in \mathcal{M}, k \in \mathcal{K}} R_1^c(h_{im}, g_{ik}). \\ \text{s.t. } R_2^c(h_{im}, g_{ik}) &\geq R_{th}. \end{aligned} \quad (12)$$

It is straightforward to see that both optimization problems **P1** and **P2** are NP-hard problems, which mean that the global optimal solutions to these two problems cannot be efficiently found. Finding the optimal combinations of antennas at both the BS and UEs may require an exhaustive search with the complexity of $\mathcal{O}(NMK)$. This becomes unaffordable when N , M and K become large. Motivated by this, in the subsequent two sections we will develop joint AS algorithms for F-NOMA and CR-NOMA systems with significantly reduced computational complexity, respectively.

III. ANTENNA SELECTION FOR F-NOMA SYSTEMS

In this section, two low-complexity joint AS algorithms are developed to maximize the system sum-rate for F-NOMA systems without and with the consideration of user fairness. The closed-form expressions for the average sum-rate of these two proposed algorithms are derived in high SNR regime.

A. Proposed AS Algorithms for F-NOMA Systems

The sum-rate R_{sum} defined in (8) is a monotonically increasing function of γ_i^s . Based on this observation, we first develop a max-max-max AS (A³-AS) scheme for the considered F-NOMA system, in which the channel gain of the instantaneous *strong* user (i.e., γ_i^s) is maximized.

1) *Max-Max-Max Antenna Selection (A³-AS)*: A³-AS mainly consists of the following three stages.

- **Stage 1.** Find out the largest elements h_i^{\max} and g_i^{\max} in each row of \mathbf{H} and \mathbf{G} , respectively. Mathematically, we have

$$h_i^{\max} = \max(h_{i1}, \dots, h_{iM}), \quad (13)$$

$$g_i^{\max} = \max(g_{i1}, \dots, g_{iK}). \quad (14)$$

for $i \in \mathcal{N}$. Each pair (h_i^{\max}, g_i^{\max}) is then treated as one AS candidate. The set of all the N pairs can be written as $\mathcal{S}_{A^3}^{(1)} = \{(h_1^{\max}, g_1^{\max}), \dots, (h_N^{\max}, g_N^{\max})\}$.

- **Stage 2.** Find out the *larger* element in each pair of (h_i^{\max}, g_i^{\max}) . Mathematically, we have

$$\gamma_i^s = \max(h_i^{\max}, g_i^{\max}), \quad i \in \mathcal{N}. \quad (15)$$

The set of the N *larger* elements are denoted by $\mathcal{S}_{A^3}^{(2)} = \{\gamma_1^s, \dots, \gamma_N^s\}$.

- **Stage 3.** Find out the largest element in $\mathcal{S}_{A^3}^{(2)}$, i.e.,

$$\gamma_{A^3}^s = \max(\gamma_i^s), \quad i \in \mathcal{N}. \quad (16)$$

It is worth noting that $\gamma_{A^3}^s$ coming from UE₁ or UE₂ is not static, but varies based on the instantaneous channel conditions of UEs and the design of AS schemes. We use (i_A^*, m_A^*) to denote the original row and column indexes of $\gamma_{A^3}^s$ when it lies in \mathbf{H} . In this case, the i_A^* th antenna at the BS and the m_A^* th antenna at UE₁ are selected, respectively. We use k_A^* to denote the original column index of $\gamma_{A^3}^w = g_{i_A^*}^{\max}$. Therefore the k_A^* th antenna at UE₂ would be selected. For the case that $\gamma_{A^3}^s$ lies in \mathbf{G} , the selected antenna indexes can be obtained in a similar way.

In A³-AS, the channel gain of the instantaneous *strong* user $\gamma_{A^3}^s$ is maximized. In this sense, the achievable rate of the instantaneous *weak* user cannot be guaranteed and could be low for some cases. In order to improve the user fairness in F-NOMA systems, we subsequently develop a max-min-max AS (AIA-AS) scheme, which aims to maximize the channel gain of the instantaneous *weak* user (i.e., γ_i^w).

2) *Max-Min-Max Antenna Selection (AIA-AS)*: Similar to the A³-AS algorithm, the AIA-AS scheme also has three stages. The main difference between A³-AS and AIA-AS lies in the second stage, where A³-AS selects the *larger* element of each pair (h_i^{\max}, g_i^{\max}) , while AIA-AS selects the *smaller* one of each pair (h_i^{\max}, g_i^{\max}) . The three stages of AIA-AS are elaborated as follow.

- **Stage 1.** Construct the set $\mathcal{S}_{AIA}^{(1)} = \{(h_1^{\max}, g_1^{\max}), \dots, (h_N^{\max}, g_N^{\max})\}$ as (13) and (14).
- **Stage 2.** Find out the *smaller* element γ_i^w in each pair (h_i^{\max}, g_i^{\max}) . That is,

$$\gamma_i^w = \min(h_i^{\max}, g_i^{\max}), \quad i \in \mathcal{N}. \quad (17)$$

The set containing these N smaller elements is denoted by $\mathcal{S}_{\text{AIA}}^{(2)} = \{\gamma_1^w, \dots, \gamma_N^w\}$.

- **Stage 3.** Find out the largest element in $\mathcal{S}_{\text{AIA}}^{(2)}$, i.e.,

$$\gamma_{\text{AIA}}^w = \max(\gamma_i^w), \quad i \in \mathcal{N}. \quad (18)$$

Similarly, we can find that γ_{AIA}^w coming from UE₁ or UE₂ is not static. We use (i_1^*, k_1^*) to denote the original row and column indexes of γ_{AIA}^w when it lies in \mathbf{G} . In this case, the i_1^* th antenna at the BS and the k_1^* th antenna at UE₂ are selected, respectively. Meanwhile, we use m_1^* to denote the original column index of $\gamma_{\text{AIA}}^s = h_{i_1^*}^{\max}$. Therefore the m_1^* th antenna at UE₁ would be selected. For the case that γ_{AIA}^w lies in \mathbf{H} , the selected antenna indexes can be obtained in a similar way.

3) *User Fairness in F-NOMA Systems:* To evaluate the user fairness of the proposed two AS algorithms in F-NOMA systems, the Jain's fairness index [24] is adopted in this paper. Specifically, the Jain's fairness index for the considered two-users scenario can be expressed as

$$\eta = \frac{(R_1 + R_2)^2}{2(R_1^2 + R_2^2)}. \quad (19)$$

We can observe that the Jain's fairness index is bounded between 0 and 1 with the maximum achieved by equaling UEs's rates.

4) *Computational Complexities of the New Algorithms for F-NOMA Systems:* As mentioned before, the complexity of the optimal selection algorithm achieved by the ES scheme in F-NOMA systems is as high as $\mathcal{O}(NMK)$. In other words, the ES-based scheme needs to calculate the achievable rates for all the NMK combinations before finding out the optimal antenna triple. When the number of antennas at each node becomes large, the computational burden would increase significantly.

The two new AS algorithms for F-NOMA systems dramatically reduce the selection complexity to $\mathcal{O}(N(M + K + 3))$, where the main computation lies in sorting the channel gains. For the case $N = M = K$, we can find that the complexities of A³-AS and AIA-AS are approximately $\mathcal{O}(N^2)$, which is an order of magnitude lower compared to that of the optimal ES-based scheme, which is $\mathcal{O}(N^3)$.

B. Performance Analysis of the New Algorithms for F-NOMA systems

In this subsection, we analyze the average sum-rate of the proposed joint AS algorithms, i.e., A³-AS and AIA-AS in F-NOMA systems, respectively.

Assuming flat Raleigh fading, $h_{im} = |\tilde{h}_{im}|^2 \geq 0$ is then an exponentially distributed random variable. When $x \geq 0$, the CDF and PDF of h_{im} are respectively given by

$$F_h(x) = 1 - e^{-\Omega_h x}, \quad (20)$$

$$f_h(x) = \Omega_h e^{-\Omega_h x}, \quad (21)$$

where $\Omega_h = 1/\mathbb{E}[h_{im}]$. Similarly, for any element in \mathbf{G} (e.g., g_{ik}), when $x \geq 0$ we have the CDF and PDF of g_{ik} as follows

$$F_g(x) = 1 - e^{-\Omega_g x}, \quad (22)$$

$$f_g(x) = \Omega_g e^{-\Omega_g x}, \quad (23)$$

where $\Omega_h = 1/\mathbb{E}[g_{ik}]$.

Recall that both the two proposed AS schemes for F-NOMA need to find the maximum element in each row of \mathbf{H} and \mathbf{G} (i.e., h_i^{\max} and g_i^{\max}). We thus first obtain the distribution functions of h_i^{\max} for $x \geq 0$ as below

$$F_{h_i^{\max}}(x) = (1 - e^{-\Omega_h x})^M \stackrel{(c_1)}{=} \sum_{i=0}^M \mu_{i,M} e^{-i\Omega_h x}, \quad (24)$$

$$f_{h_i^{\max}}(x) = -\sum_{i=1}^M i\Omega_h \mu_{i,M} e^{-i\Omega_h x}, \quad (25)$$

where $\mu_{i,M} = (-1)^i \binom{M}{i}$ and the expansion step (c_1) is conducted based on the Binomial theorem.

Similarly, the CDF and PDF of g_i^{\max} for $x \geq 0$ are given by

$$F_{g_i^{\max}}(x) = (1 - e^{-\Omega_g x})^K = \sum_{j=0}^K \mu_{j,K} e^{-j\Omega_g x}, \quad (26)$$

$$f_{g_i^{\max}}(x) = -\sum_{j=1}^K j\Omega_g \mu_{j,K} e^{-j\Omega_g x}. \quad (27)$$

From the sum-rate expression of the considered F-NOMA system in (8), we can observe that the achievable rate of the instantaneous *weak* UE approaches a constant when the transmit SNR $\rho \rightarrow \infty$, i.e., $\log_2 \left(1 + \frac{a\gamma_i^w}{b\gamma_i^w + 1/\rho}\right) \stackrel{\rho \rightarrow \infty}{\approx} \log_2(1/b)$, which is only related to the power allocation coefficient b . In this case, when the power allocation coefficient is fixed in F-NOMA systems, we only need to focus on the contribution of the instantaneous *strong* UE, i.e., γ_i^s , to the sum-rate. Mathematically, we have the approximation of the system instantaneous sum-rate given in (8) as follows:

$$R_{\text{sum}}(\gamma_i^s) \stackrel{\rho \rightarrow \infty}{\approx} \log_2(1 + b\rho\gamma_i^s) + \log_2 \frac{1}{b}. \quad (28)$$

It is worth pointing out that γ_i^s coming from UE₁ or UE₂ is not only determined by their geographical locations, but also by the small scale fading coefficient. Subsequently, we perform the asymptotic analysis for the average sum-rates achieved by A³-AS and AIA-AS algorithms, respectively.

1) Analytical Average Sum-Rate of the A³-AS Algorithm: Let $h^{\max} = \max(h_{ik})$ and $g^{\max} = \max(g_{ik})$ for $i \in \mathcal{N}$, $m \in \mathcal{M}$ and $k \in \mathcal{K}$. Recall that in A³-AS, $\gamma_{A^3}^s$ is actually the relatively larger element of h^{\max} and g^{\max} . In order to obtain the distribution of $\gamma_{A^3}^s$, we first respectively derive the CDF and PDF of h^{\max}

and g^{\max} given by

$$F_{h^{\max}}(x) = \sum_{i=0}^{NM} \mu_{i,NM} e^{-i\Omega_h x}, \quad f_{h^{\max}}(x) = -\sum_{i=1}^{NM} i\Omega_h \mu_{i,NM} e^{-i\Omega_h x}, \quad (29)$$

$$F_{g^{\max}}(x) = \sum_{j=0}^{NK} \mu_{j,NK} e^{-j\Omega_g x}, \quad f_{g^{\max}}(x) = -\sum_{j=1}^{NK} j\Omega_g \mu_{j,NK} e^{-j\Omega_g x}. \quad (30)$$

Then the CDF and PDF of $\gamma_{A^3}^s = \max(h^{\max}, g^{\max})$ in A³-AS can be calculated as follows

$$F_{\gamma_{A^3}^s}(x) = \Pr\{\max(h^{\max}, g^{\max}) < x\} = \Pr(h^{\max} \leq g^{\max} < x) + \Pr(g^{\max} < h^{\max} < x), \quad (31)$$

$$\begin{aligned} f_{\gamma_{A^3}^s}(x) &= f_{g^{\max}}(x) \int_0^x f_{h^{\max}}(y) dy + f_{h^{\max}}(x) \int_0^x f_{g^{\max}}(y) dy \\ &= \sum_{i=1}^{NM} \sum_{j=1}^{NK} \mu_{i,NM} \mu_{j,NK} (i\Omega_h e^{-i\Omega_h x} + j\Omega_g e^{-j\Omega_g x} - (i\Omega_h + j\Omega_g) e^{-(i\Omega_h + j\Omega_g)x}). \end{aligned} \quad (32)$$

We now can analyze the average sum-rate of A³-AS for high SNR scenarios and obtain Proposition 1 as below.

Proposition 1. *With the PDF of $\gamma_{A^3}^s$ in A³-AS, i.e., $f_{\gamma_{A^3}^s}(x)$, when $\rho \rightarrow \infty$, the asymptotic closed-form expression for the average sum-rate of the A³-AS algorithm is given by*

$$\bar{R}_{\text{sum}}^{A^3} \approx \frac{1}{\ln 2} (\ln \rho - C) + \frac{1}{\ln 2} \sum_{i=1}^{NM} \sum_{j=1}^{NK} \mu_{i,NM} \mu_{j,NK} \ln \left(\frac{i\Omega_h + j\Omega_g}{ij\Omega_h \Omega_g} \right), \quad (33)$$

Proof: See Appendix A. ■

As can be observed from (33), the system sum-rate $\bar{R}_{\text{sum}}^{A^3}$ is an increasing function of the transmit SNR ρ as expected. Interestingly, it is not affected by the power allocation coefficient b . The reason behind this observation is that when $\rho \rightarrow \infty$, $R_{\text{sum}} \approx \log_2(\gamma^s \rho)$, in which the power allocation factor b has been removed. It will be verified via the computer simulations in section V.

2) *Analytical Average Sum-Rate of AIA-AS Algorithm:* In this subsection, we derive the asymptotic average sum-rate of the proposed AIA-AS algorithm, i.e., $\bar{R}_{\text{sum}}^{\text{AIA}}$, for high SNR scenarios. According to (28), we first need to obtain the PDF of the channel gain of the instantaneous *strong* UE, $f_{\gamma_{\text{AIA}}^s}(x)$, which is provided by Lemma 1 as below.

Lemma 1. *In AIA-AS algorithm, the PDF of the instantaneous channel gain of the strong user, i.e., γ_{AIA}^s , is given by*

$$f_{\gamma_{\text{AIA}}^s}(x) = \sum_{i=1}^M \sum_{j=1}^K \sum_{\ell} C_{\ell} t_{\ell} \zeta_{ij} (\psi(i\Omega_h, j\Omega_g) + \psi(j\Omega_g, i\Omega_h)), \quad (34)$$

where the multinomial coefficient $C_{\ell} = \frac{(N-1)!}{\ell_0! \dots \ell_{MK}!}$ for $\ell_0 + \dots + \ell_{MK} = N-1$, $t_{\ell} = \prod_{1 \leq i \leq M, 1 \leq j \leq K} (-\mu_{i,M} \mu_{j,K})^{\ell_{ij}}$, $\zeta_{ij} = N i j \Omega_h \Omega_g \mu_{i,M} \mu_{j,K}$, $\xi_{\ell} = \sum_{i=1}^M \sum_{j=1}^K (i\Omega_h + j\Omega_g) \ell_{ij}$, and $\psi(\theta_1, \theta_2) = e^{-\theta_1 x} \left(\frac{e^{-\theta_2 x} - 1}{\theta_2} - \frac{e^{-(\theta_2 + \xi_{\ell})x} - 1}{\theta_2 + \xi_{\ell}} \right)$.

Proof: See Appendix B. ■

Provided the PDF of γ_{AIA}^s in AIA-AS algorithm, we can approximate the average sum-rate achieved by AIA-AS in Proposition 2 as below.

Proposition 2. *With the PDF of γ_{AIA}^s , when $\rho \rightarrow \infty$, the asymptotic average sum-rate of AIA-AS, i.e., $\bar{R}_{\text{sum}}^{\text{AIA}}$, is given by*

$$\bar{R}_{\text{sum}}^{\text{AIA}} \approx \log_2 \frac{1}{b} + \sum_{i=1}^M \sum_{j=1}^K \sum_{\ell} \frac{C_{\ell} t_{\ell}}{\ln 2} (T_1 + T_2 + T_3 + T_4), \quad (35)$$

where T_n , $n = \{1, 2, 3, 4\}$ are given by

$$\begin{aligned} T_1 &= \frac{\xi_{\ell} \tilde{\zeta}_{ij}}{\phi_i} \chi(j\Omega_g), & T_2 &= \frac{\xi_{\ell} \tilde{\zeta}_{ij}}{\phi_j} \chi(i\Omega_h), \\ T_3 &= \frac{\zeta_{ij} \phi_{ij,2} \chi(\phi_{ij,1})}{\phi_i \phi_j \phi_{ij,1}}, & T_4 &= -\tilde{\zeta}_{ij} \chi(i\Omega_h + j\Omega_g), \end{aligned}$$

in which, $\tilde{\zeta}_{ij} = N\mu_{i,M}\mu_{j,K}$, $\phi_i = i\Omega_h + \xi_{\ell}$, $\phi_j = j\Omega_g + \xi_{\ell}$, $\phi_{ij,1} = i\Omega_h + j\Omega_g + \xi_{\ell}$, $\phi_{ij,2} = i\Omega_h + j\Omega_g + 2\xi_{\ell}$, $\chi(x) = C + \ln \frac{x}{b\rho}$, and the constant C is the Euler's constant.

Proof: See Appendix C. ■

IV. ANTENNA SELECTION FOR CR-NOMA SYSTEMS

In this section, CR-NOMA is considered, and a computationally efficient joint AS algorithm, i.e., MCG-AS, is proposed to maximize the achievable rate of secondary user UE₁. To further reduce the computational complexity, two simplified versions of MCG-AS, i.e., PU-AS and SU-AS, are proposed. In the high SNR regime, the asymptotic expressions of average achievable rates of UE₁ using MCG-AS, PU-AS and SU-AS are then derived, respectively.

A. Proposed AS Algorithms for CR-NOMA Systems

By observing (9)-(11), we can see that the channel gain of the primary user UE₂ (i.e., g_{ik}) and that of the secondary user UE₁ (i.e., h_{im}) contribute to R_1^c in different ways in CR-NOMA systems. In particular, if g_{ik} increases, less power would be needed to satisfy the QoS requirement of UE₂. In other words, more power can be allocated to UE₁ and hence R_1^c increases. In contrast, the contribution of h_{im} to R_1^c does not lie in the power domain, but can boost R_1^c directly via providing a better channel gain.

Based on the above observation, we develop a low-complexity AS algorithm, referred to as maximum-channel-gain-based antenna selection (MCG-AS), for the considered CR-NOMA system. The key idea of

MCG-AS is to jointly select the antenna triple from all the nodes by maximizing the maximum channel gain from the two UEs. More specifically, the MCG-AS algorithm consists of four stages as below.

- **Stage 1.** Find out the largest element h^{\max} from \mathbf{H} and g^{\max} from \mathbf{G} .

$$h^{\max} = \max(h_{i1}, \dots, h_{NM}), \quad (36)$$

$$g^{\max} = \max(g_{i1}, \dots, g_{NK}). \quad (37)$$

- **Stage 2.** Compare and find out the larger element between h^{\max} and g^{\max} . That is,

$$\gamma_{\text{MCG}}^s = \max(h^{\max}, g^{\max}). \quad (38)$$

Denote the row index of γ_{MCG}^s as i_{MCG}^* . Then, the i_{MCG}^* th antenna at the BS is selected.

- **Stage 3.** Find out γ_{MCG}^w which lies in the same row of γ_{MCG}^s . Mathematically, we have

$$\gamma_{\text{MCG}}^w = \begin{cases} \max(g_{i_{\text{MCG}}^*1}, \dots, g_{i_{\text{MCG}}^*K}), & \text{if } \gamma_{\text{MCG}}^s = h^{\max}, \\ \max(h_{i_{\text{MCG}}^*1}, \dots, h_{i_{\text{MCG}}^*M}), & \text{if } \gamma_{\text{MCG}}^s = g^{\max}. \end{cases} \quad (39)$$

For the case that $\gamma_{\text{MCG}}^s = h^{\max}$, we denote the original column indexes of γ_{MCG}^s and γ_{MCG}^w as m_{MCG}^* and k_{MCG}^* , respectively. Then, the m_{MCG}^* th antenna at UE₁ and the k_{MCG}^* th antenna at UE₂ are selected concurrently. For the case that $\gamma_{\text{MCG}}^s = g^{\max}$, the corresponding antenna indexes at UEs can be selected similarly.

- **Stage 4.** Obtain b^c and a^c by substituting γ_{MCG}^s and γ_{MCG}^w to (10).

We also realize that, when the primary user UE₂ is much closer to the BS than the secondary user UE₁, it is likely that g^{\max} is larger than h^{\max} . In this case, MCG-AS can be simplified into the primary-user-based AS (PU-AS) scheme, in which the AS can be performed by maximizing the channel gain of the primary user UE₂. In this sense, more power can be allocated to the secondary user UE₁. In contrast, when UE₂ is much farther from the BS than UE₁, it is likely that g^{\max} is smaller than h^{\max} . In this case, MAG-AS can be reduced to the secondary-user-based AS (SU-AS) scheme, which aims to maximize the channel gain of the secondary user UE₁ directly. We elaborate the principles and key stages of the simplified PU-AS and SU-AS in the subsequent two subsections.

1) Primary-User-based Antenna Selection (PU-AS): The key idea of PU-AS is to select the antenna pair from the BS and the primary user UE₂ concurrently, which provides the maximum channel gain for UE₂. In this case, less power is needed for UE₂ to satisfy its QoS requirement and more power can thus be allocated to the secondary user UE₁ to increase the corresponding R_1^c . The main three stages of PU-AS scheme are described as below.

- **Stage 1.** Find out the largest element g^{\max} from \mathbf{G} .

$$g^{\max} = \max(g_{i1}, \dots, g_{NK}). \quad (40)$$

Denote the row and column indexes of g^{\max} as i_p^* and k_p^* , respectively. Then, the i_p^* th antenna at the BS and the k_p^* th antenna at UE₂ are selected simultaneously.

- **Stage 2.** Find out the largest element $h_{i_p^*}^{\max}$ in \mathbf{H} . That is,

$$h_{i_p^*}^{\max} = \max(h_{i_p^*1}, \dots, h_{i_p^*M}). \quad (41)$$

Denote the column index of $h_{i_p^*}^{\max}$ as m_p^* . Then, the m_p^* th antenna at UE₁ is selected.

- **Stage 3.** Calculate b^c and a^c by substituting the value of g^{\max} and $h_{i_p^*}^{\max}$ to (10).

2) *Secondary-User-based Antenna Selection (SU-AS)*: The key idea of SU-AS is to first select the antenna pair from the BS and secondary user UE₁ concurrently, which provides the maximum channel gain for UE₁. Similarly, SU-AS consists of three stages.

- **Stage 1.** Find out the largest element h^{\max} from \mathbf{H} .

$$h^{\max} = \max(h_{i1}, \dots, h_{NM}). \quad (42)$$

Denote the row and column indexes of h^{\max} as i_s^* and m_s^* , respectively. Then, the i_s^* th antenna at the BS and the m_s^* th antenna at UE₁ are selected simultaneously.

- **Stage 2.** Find out the largest element $g_{i_s^*}^{\max}$ in \mathbf{G} . Mathematically,

$$g_{i_s^*}^{\max} = \max(g_{i_s^*1}, \dots, g_{i_s^*K}). \quad (43)$$

Denote the column index of $g_{i_s^*}^{\max}$ as k_s^* . Then, the k_s^* th antenna at UE₂ is selected.

- **Stage 3.** Obtain b^c and a^c by substituting the value of h^{\max} and $g_{i_s^*}^{\max}$ to (10).

3) *Computational Complexities of Proposed Algorithms for CR-NOMA Systems*: Similar to F-NOMA systems, the complexity of the optimal selection algorithm achieved by the ES-based scheme in CR-NOMA systems is also as high as $\mathcal{O}(NMK)$. This is because the ES-based scheme needs to calculate the achievable rates of the secondary user UE₁ for all the NMK combinations before finding out the optimal antenna triple. When the number of antennas at each node becomes large, the computational burden would increase significantly.

In contrast, all the three proposed AS algorithms, i.e., MCG-AS, PU-AS and SU-AS, can dramatically reduce the selection complexities as illustrated in Table I, where the main computation lies in sorting the corresponding channel gains. Furthermore, the complexities of PU-AS and SU-AS are lower than that of MCG-AS. For the case $N = M = K$, the computational burdens of PU-AS and SU-AS are nearly half of that of MCG-AS.

TABLE I: Complexity analysis for different antenna selection algorithms in CR-NOMA systems

CR-NOMA	Computational Complexity			
	CR-NOMA-ES	MCG-AS	PU-AS	SU-AS
	$\mathcal{O}(NMK)$	$\mathcal{O}(N(M+K)+2)$	$\mathcal{O}(NK+M)$	$\mathcal{O}(NM+K)$

B. Analytical Average Rate of the Secondary User in CR-NOMA Systems

As discussed above, since the QoS of the primary user UE_2 is guaranteed in CR-NOMA systems, we thus only focus on the analysis of the average rate of the secondary user UE_1 , which is served opportunistically in CR-NOMA networks.

For the proposed MCG-AS scheme, by carefully analyzing its antenna selection procedure, we can find that the MCG-AS scheme would become SU-AS when $h^{\max} \geq g^{\max}$ and would simplify to PU-AS otherwise. In this case, the average rate of UE_1 in MCG-AS (i.e., \bar{R}_1^{MCG}) can be calculated according to the law of total probability as below:

$$\bar{R}_1^{\text{MCG}} = \Pr(h^{\max} < g^{\max}) \bar{R}_1^{\text{P}} + \Pr(h^{\max} \geq g^{\max}) \bar{R}_1^{\text{S}}, \quad (44)$$

where \bar{R}_1^{P} and \bar{R}_1^{S} denote the average rate of UE_1 in PU-AS and SU-AS, respectively, and the probability of the event that $h^{\max} \geq g^{\max}$ is given by

$$\Pr(h^{\max} \geq g^{\max}) = \int_0^\infty f_{h^{\max}}(x) \int_0^x f_{g^{\max}}(y) dy dx = \sum_{i=1}^{NM} \sum_{j=1}^{NK} \nu_{ij} \mu_{i,NM} \mu_{j,NK}, \quad (45)$$

where $\nu_{ij} = j\Omega_g / (i\Omega_h + j\Omega_g)$. Accordingly, $\Pr(h^{\max} < g^{\max}) = 1 - \Pr(h^{\max} \geq g^{\max})$. Therefore, the analytical expression of \bar{R}_1^{MCG} can be obtained if we can respectively derive the expressions of \bar{R}_1^{P} and \bar{R}_1^{S} .

We first address the calculation of \bar{R}_1^{P} . Recall that in the PU-AS scheme, the maximum element of \mathbf{G} , i.e., g^{\max} , is first selected, and then the maximum element in the corresponding row of \mathbf{H} is subsequently selected. By using the i.i.d. property of the all the elements in \mathbf{H} , the PDF of $h_{i_p^*}^{\max}$ can be derived by using $f_{h_i^{\max}}(x)$. In this case, by given the PDF of g^{\max} in (30) and that of $h_{i_p^*}^{\max}$ in (25), an approximated expression of \bar{R}_1^{P} can be given in Proposition 3 as below.

Proposition 3. *With the QoS requirement of the primary user R_{th} and the distributions of g^{\max} and $h_{i_p^*}^{\max}$, the average rate of the secondary user UE_1 in PU-AS, i.e., \bar{R}_1^{P} , can be approximated in the high SNR*

regime as below:

$$\bar{R}_1^P \approx \sum_{i=1}^M \sum_{j=1}^{NK} \frac{\mu_{i,M} \mu_{i,NK}}{\ln 2} \left(\frac{\varepsilon j \Omega_g}{i \Omega_h - \varepsilon j \Omega_g} \ln \frac{(\varepsilon + 1) j \Omega_g}{i \Omega_h + j \Omega_g} - \ln \frac{i \Omega_h}{\rho} - C \right). \quad (46)$$

Proof: See Appendix D. ■

After obtaining \bar{R}_1^P , we thus turn to calculate \bar{R}_1^S for SU-AS. Recall that in SU-AS, the maximum element of \mathbf{H} (i.e., h^{\max}) is selected first, followed by the corresponding maximum element in the i_S^* row of \mathbf{G} (i.e., $g_{i_S^*}^{\max}$). Similarly, considering the i.i.d. property of the all the elements in \mathbf{G} , the PDF of $g_{i_S^*}^{\max}$ can be characterized by $f_{g_i^{\max}}(x)$. Given the PDF of h^{\max} in (29) and that of $g_{i_S^*}^{\max}$ in (27), the average rate of UE₁ in SU-AS, i.e., \bar{R}_1^S , is given by Proposition 4 as below.

Proposition 4. *With the QoS requirement of the primary UE and the distributions of h^{\max} and $g_{i_S^*}^{\max}$, respectively, the average rate of UE₁ in SU-AS, i.e., \bar{R}_1^S , can be approximated in the high SNR regime as below:*

$$\bar{R}_1^S \approx \sum_{i=1}^{NM} \sum_{j=1}^K \frac{\mu_{i,M} \mu_{i,NK}}{\ln 2} \left(\frac{\varepsilon j \Omega_g}{i \Omega_h - \varepsilon j \Omega_g} \ln \frac{(\varepsilon + 1) j \Omega_g}{i \Omega_h + j \Omega_g} - \ln \frac{i \Omega_h}{\rho} - C \right). \quad (47)$$

Proof: Similar to Proposition 3, \bar{R}_1^S can be obtained by resolving the integral (11) over h^{\max} and $g_{i_S^*}^{\max}$. Comparing the distributions of h^{\max} and $g_{i_S^*}^{\max}$ in Proposition 4 to those of $h_{i_p^*}^{\max}$ and g^{\max} in Proposition 3, we can observe that the only difference between them lies in the corresponding upper limits of the summations. In this case, we can calculate \bar{R}_1^S by replacing the upper limits of i and j in (46) by NM and K , respectively. ■

Finally, by substituting (45)-(47) into (44), we can obtain an approximated closed-form expression for \bar{R}_1^{MCG} .

Remark 1. *When the primary user UE₂ is much closer to the BS than the secondary user UE₁, the event that $g^{\max} > h^{\max}$ occurs with a large probability. In this case, the MCG-AS scheme will reduce to the PU-AS scheme and thus $\bar{R}_1^{\text{MCG}} \approx \bar{R}_1^P$. On the other hand, when UE₂ is much further from the BS than UE₁, it is more likely that $g^{\max} \leq h^{\max}$. In this case, the MCG-AS scheme degrades to SU-AS, and we have the approximation $\bar{R}_1^{\text{MCG}} \approx \bar{R}_1^S$.*

V. NUMERICAL PERFORMANCE EVALUATIONS

In this section, the performance of the proposed joint AS algorithms for both F-NOMA and CR-NOMA systems is evaluated by using computer simulations. In all simulations, we set $M = K = 2$, $\Omega_h = d_1^\alpha$,

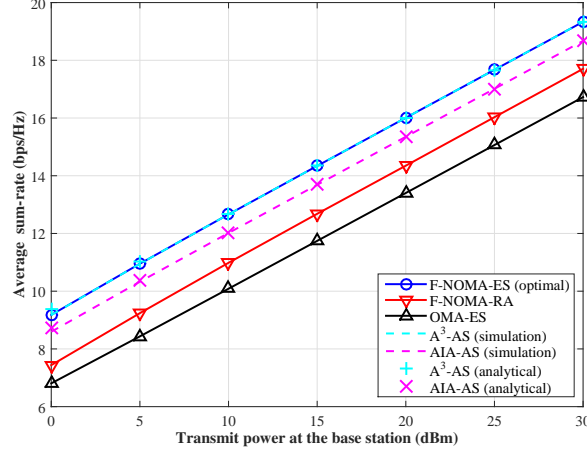


Fig. 2: Average sum-rate vs. transmit power in F-NOMA, $N = 2$, $d_1 = 30\text{m}$, $d_2 = 100\text{m}$, $a = 0.6$, $b = 0.4$, $\sigma^2 = -70\text{dBm}$.

$\Omega_g = d_2^\alpha$, d_1 (d_2) is the distance between the BS and UE_1 (UE_2) and the path-loss exponent is set as $\alpha = 3$.

A. Numerical Results for F-NOMA Systems

Fig. 2 illustrates how the transmit power P_s at the BS affects the system average sum-rate \bar{R}_{sum} . As can be observed from Fig. 2, when P_s increases, \bar{R}_{sum} increases for all the schemes. Moreover, the performance of the proposed $\text{A}^3\text{-AS}$ and AIA-AS schemes is much better than that of the random AS scheme in F-NOMA scenarios (F-NOMA-RA), since both $\text{A}^3\text{-AS}$ and AIA-AS utilize the spatial degrees of freedom brought by the multiple antennas at each node. Furthermore, the $\text{A}^3\text{-AS}$ scheme can achieve the same performance as that of the optimal ES scheme in NOMA scenarios (F-NOMA-ES) but with a much lower computational complexity. We should note that the analytical results match the simulation results for both $\text{A}^3\text{-AS}$ and AIA-AS , which validate our theoretical analysis in Sec. III. It is also worth pointing out that all the NOMA schemes outperform the ES scheme in OMA systems (OMA-ES) over the entire SNR region. For simplicity, we only show the analytical results in the following discussions.

Fig. 3 illustrates how the number of antennas N at the BS influences the average sum-rate \bar{R}_{sum} . We can see from this figure that the sum-rates of F-NOMA-RA and AIA-AS keep constant when N increases. For the F-NOMA-RA scheme, this is because it does not properly utilize the multiple-antenna setting but selects one antenna at each node randomly. The reason why AIA-AS keeps constant is because it guarantees the performance of the *weak* user with the poor channel gain γ^w , but not the *strong* user with

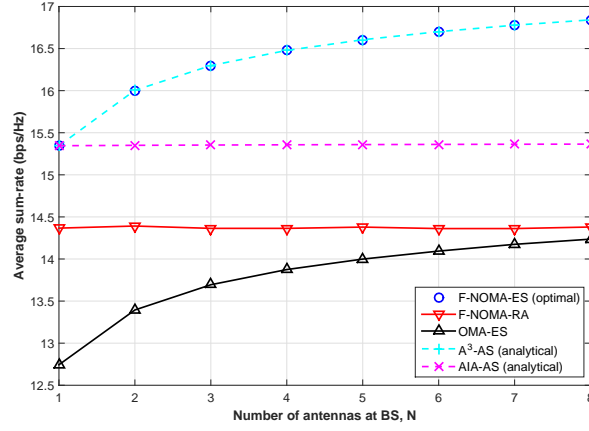


Fig. 3: Average sum-rate vs. N in F-NOMA, $d_1 = 30\text{m}$, $d_2 = 100\text{m}$, $a = 0.6$, $b = 0.4$, $\sigma^2 = -70\text{dBm}$, $P_s = 20\text{dBm}$.

the better channel condition γ^s , which contributes the most to \bar{R}_{sum} . In contrast, the average sum-rate of A³-AS increases along with N , and A³-AS achieves the same performance as that of the optimal F-NOMA-ES scheme. Again, all the NOMA schemes outperform the OMA-ES scheme in the entire region.

Fig. 4 depicts how the distances between the BS and users influence \bar{R}_{sum} for the schemes in F-NOMA systems. Take a constant d_1 and a variable d_2 for example. We can observe that when d_2 increases, \bar{R}_{sum} decreases for all the schemes in F-NOMA systems. We also note that both A³-AS and AIA-AS outperform the F-NOMA-RA and the OMA-ES schemes, and again A³-AS achieves almost the same performance as F-NOMA-ES. There is a crossing between the curves for F-NOMA-RA and OMA-ES. The reason for this is that in OMA-ES, when d_2 is much larger than d_1 , the energy and frequency resources allocated to UE₂ are wasted since they contribute very little to \bar{R}_{sum} .

Fig. 5 demonstrates how the power allocation coefficient b affects the \bar{R}_{sum} for various AS schemes in F-NOMA systems. Interestingly we can see that for all the F-NOMA schemes the average sum-rate keeps almost constant when b increases. The main reason is that $R_{\text{sum}} \approx \log_2(\gamma^s \rho)$ when $\rho \rightarrow \infty$; that is, the sum-rate is not affected by the value of b . In contrast, the performance of the OMA-ES scheme decreases when b increases as more power is allocated to the *weak* user with the poor channel condition γ^w which contributes little to \bar{R}_{sum} .

Although the system sum-rate performance of A³-AS is shown to be slightly better than that of AIA-AS, regarding the fairness between UE₁ and UE₂, we can observe in Fig. 6 that AIA-AS can provide better

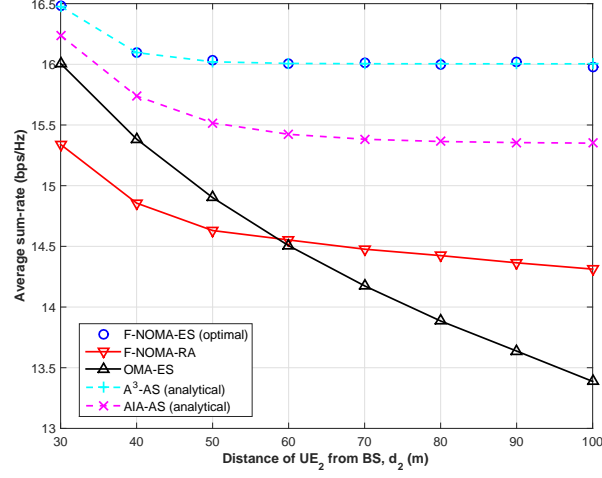


Fig. 4: Average sum-rate vs. d_2 in F-NOMA, $N = 2$, $d_1 = 30\text{m}$, $a = 0.6$, $b = 0.4$, $\sigma^2 = -70\text{dBm}$, $P_s = 20\text{dBm}$.

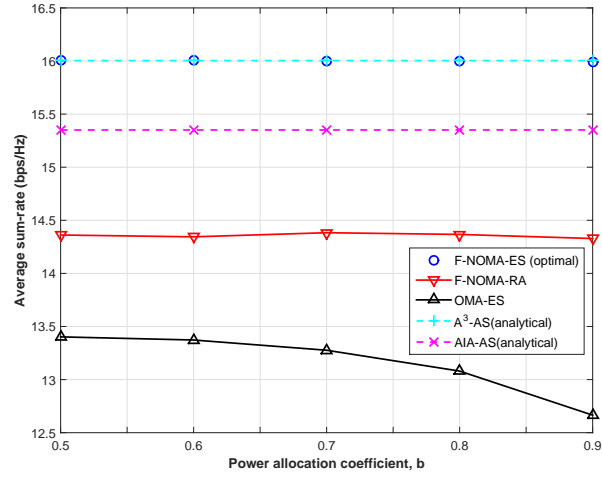


Fig. 5: Average sum-rate vs. b in F-NOMA, $N = 2$, $d_1 = 30\text{m}$, $d_2 = 100\text{m}$, $a = 1 - b$, $\sigma^2 = -70\text{dBm}$, $P_s = 20\text{dBm}$.

user fairness than A³-AS. In other words, in practice AIA-AS would be a better choice to balance the tradeoff between the system throughput and user fairness.

B. Numerical Results for CR-NOMA Systems

In the considered two-user CR-NOMA down-link scenario, without loss of generality, we treat UE₂ as the primary user and UE₁ as the secondary user. Fig. 7 illustrates how the locations of UEs affect

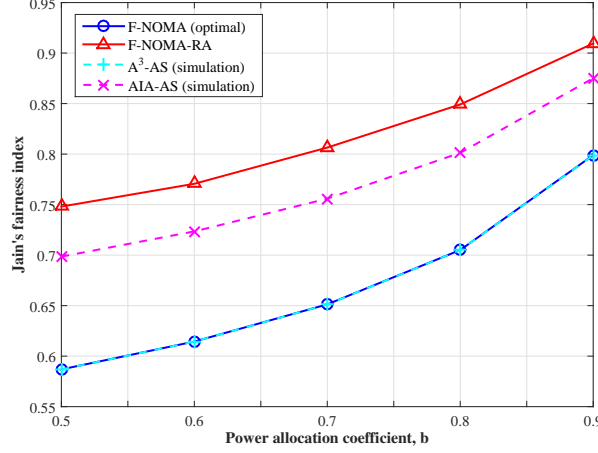


Fig. 6: Jain's fairness index vs. b in F-NOMA, $N = 4, d_1 = 60\text{m}, d_2 = 100\text{m}, a = 1 - b, \sigma^2 = -70\text{dBm}, P_s = 20\text{dBm}$.

the average rate of UE_1 . We now fix the location of the primary user UE_2 ($d_2 = 100\text{m}$) and place the secondary user UE_1 at various locations. As can be observed from Fig. 7, when d_1 increases, i.e., the distance of UE_1 from the BS goes up, the average rate of UE_1 , \bar{R}_1^c , decreases for all the proposed AS schemes in CR-NOMA systems. Interestingly, there is a crossing for the curves of PU-AS and SU-AS around $d_1 = d_2 = 100\text{m}$. \bar{R}_1^c achieved by SU-AS is larger than that of PU-AS when $d_1 \leq 100\text{m}$ and the situation is reverse when $d_1 > 100\text{m}$. The reason is that when UE_1 is closer to the BS, the channel gain of UE_1 contributes more to \bar{R}_1^c , but when UE_2 is closer to the BS, the channel gain of UE_2 is more dominant to \bar{R}_1^c as more power is allocated to the secondary UE. Furthermore, MCG-AS, which takes advantage of PU-AS and SU-AS, can achieve almost the same performance as the optimal ES scheme in CR-NOMA scenarios (CR-NOMA-ES) for all the location settings but with much lower computational complexity, and all the proposed schemes outperform the random selection scheme in CR-NOMA scenarios (CR-NOMA-RA). It is worth pointing out that the analytical results match the simulation results for all the proposed CR-NOMA AS schemes, which validates our theoretical analysis in Sec. IV. For brevity, we only show the analytical results in the following discussions.

Fig. 8 illustrates how the transmit power at the BS influences the achievable \bar{R}_1^c . We can see that when P_s increases, \bar{R}_1^c increases for all the schemes in CR-NOMA systems. Moreover, all the proposed AS schemes outperform CR-NOMA-RA. In particular, when UE_1 is closer to the BS, the performance of SU-AS is better than that of PU-AS. In contrast, when UE_2 is closer to the BS, the performances of SU-AS and PU-AS are inverted. Again MCG-AS achieves the near-optimal performance as CR-NOMA-ES in the

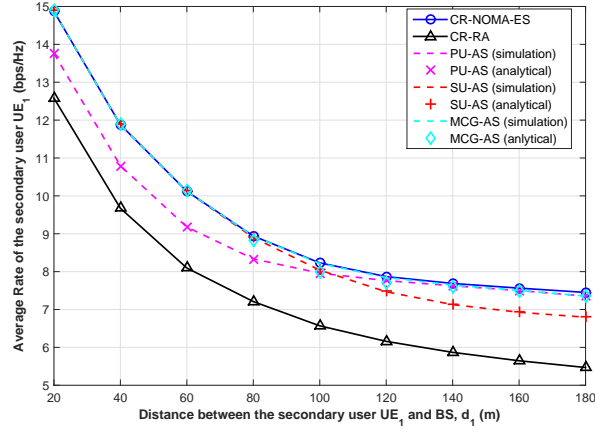


Fig. 7: Average rate of UE_1 vs. d_1 in CR-NOMA, $N = 4, d_2 = 100\text{m}, R_{th} = 5\text{bps/Hz}, \sigma^2 = -70\text{dBm}, P_s = 25\text{dBm}$.

entire region.

Fig. 9 illustrates how the number of antennas at the BS affects the achievable \bar{R}_1^c . When UE_1 is closer to the BS, the curve for SU-AS increases along with N . This is because in this case, the channel gain of UE_1 contributes more to \bar{R}_1^c and h_{ik} is maximized in SU-AS. In contrast, the curve for PU-AS keeps constant even when N increases. The reason for this is that PU-AS aims at maximizing the channel gain of UE_2 instead of UE_1 . That is, the channel gain of UE_1 is selected in a similar way to finding out the maximum element from a random row of \mathbf{H} and h_{ik} is affected by M but not N . Interestingly, when UE_1 is farther from the BS than UE_2 , the aforementioned situation is inverted. Furthermore, all the proposed schemes outperform CR-NOMA-RA, and MCG-AS can achieve a near-optimal performance in the entire region.

Fig. 10 demonstrates how the QoS requirement R_{th} affects the achievable \bar{R}_1^c . We can see that when R_{th} increases, the performances decrease for all the AS schemes in CR-NOMA scenarios. This is because when R_{th} increases, more power will be allocated to UE_2 to meet the higher QoS requirement and less power can be allocated to UE_1 . In this case, \bar{R}_1^c would decrease when R_{th} increases. Again, all the proposed AS schemes outperform CR-NOMA-RA, and MCG-AS achieves a near-optimal performance in the entire region. Moreover, the performance of PU-AS is better than that of SU-AS when UE_2 is closer to the BS, and it is opposite when UE_1 is closer to the BS.

As discussed above, MCG-AS can provide a near-optimal average rate performance for the secondary user in CR-NOMA networks by taking advantage of the instantaneous channel conditions of all the users in systems. However, once the network-accessing priority and the near-far relationship of the users in

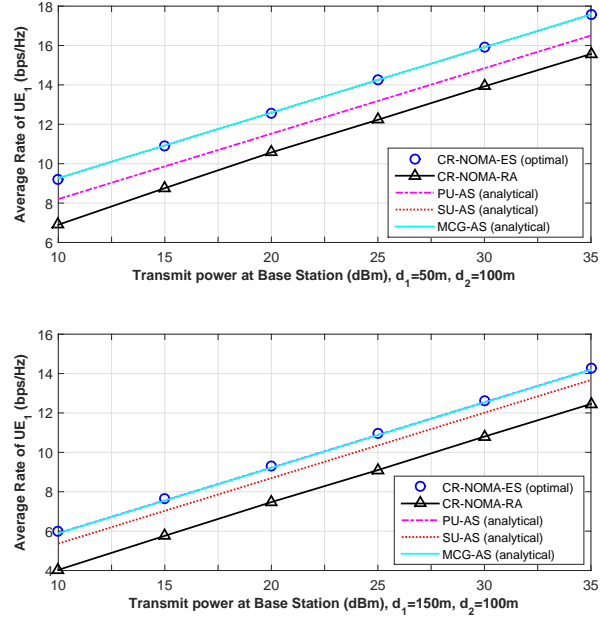


Fig. 8: Average rate of UE₁ vs. P_s in CR-NOMA, $N = 4$, $R_{\text{th}} = 5\text{bps/Hz}$, $\sigma^2 = -80\text{dBm}$.

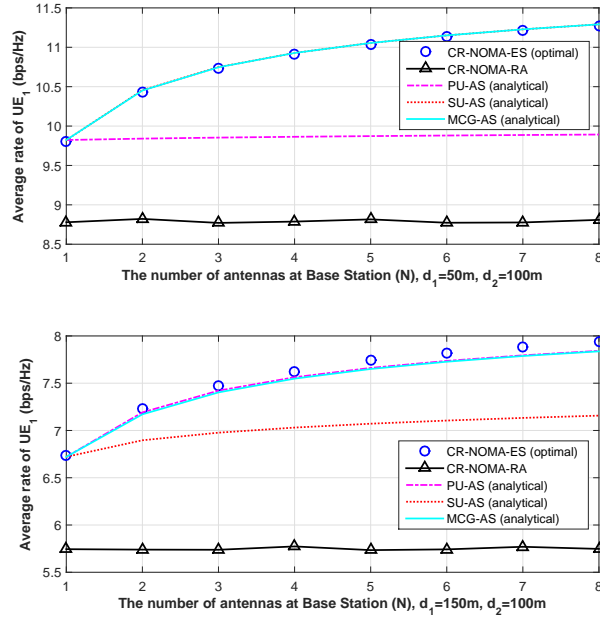


Fig. 9: Average rate of UE₁ vs. N in CR-NOMA, $P_s = 25\text{dBm}$, $R_{\text{th}} = 5\text{bps/Hz}$, $\sigma^2 = -70\text{dBm}$.

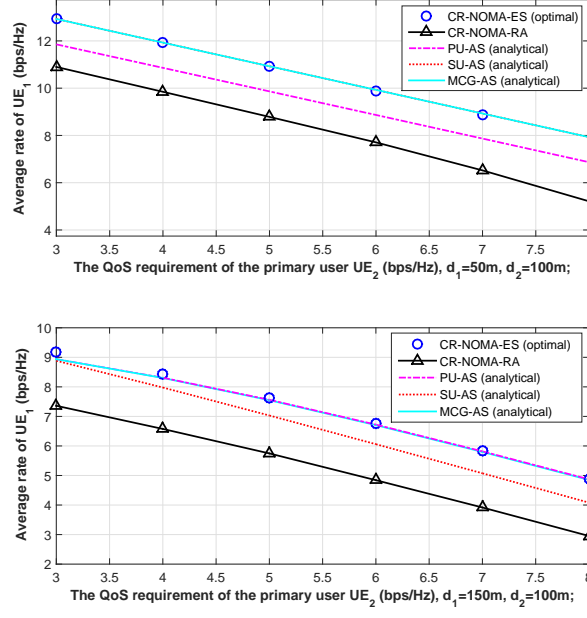


Fig. 10: Average rate of UE₁ vs. the QoS requirement of UE₂ in CR-NOMA, $N = 4$, $P_s = 25\text{dBm}$, $R_{\text{th}} = 5\text{bps/Hz}$, $\sigma^2 = -70\text{dBm}$.

CR-networks are determined, the PU-AS or SU-AS algorithms would be better choices. This is because the computational complexities of PU-AS and SU-AS are further lower than that of MCG-AS, and a near-optimal average rate performance for the secondary UE can be achieved by PU-AU when the primary user is much closer to the BS, and that is achieved by SU-AS when the secondary user is much closer to the BS.

VI. CONCLUSIONS

This paper studied the joint antenna selection (AS) problem for a classical two-user MIMO-NOMA system. Two computationally efficient AS algorithms, named A³-AS and AIA-AS, were proposed to maximize the system sum-rate of F-NOMA systems without and with the consideration of user fairness, respectively. Meanwhile, a low-complexity AS algorithm (i.e., MCG-AS) and its two simplified versions (i.e., PU-AS and SU-AS) were developed to maximize the rate of the secondary user in CR-NOMA systems. The asymptotic closed-form expressions for the average sum-rates in F-NOMA systems and average rates of the secondary user in CR-NOMA systems were provided. Numerical simulations demonstrated that in F-NOMA systems, both A³-AS and AIA-AS yield significant performance gains over the random AS scheme and the OMA scheme with exhaustive search-based AS. Furthermore, A³-AS achieves the

near-optimal sum-rate performance while AIA-AS provides better user fairness. In CR-NOMA scenarios, MCG-AS can achieve a near-optimal performance with much reduced complexity in the entire region. The complexities of the PU-AS and SU-AS is further lower than that of the MCG-AS scheme, while they only achieve a near-optimal performance when one user is much closer to the base station than the other.

APPENDIX A

PROOF OF THE PROPOSITION 1

When $\rho \rightarrow \infty$, we can attain the asymptotic closed-form expression for the average sum-rate of the A³-AS algorithm as follows:

$$\begin{aligned}
\bar{R}_{\text{sum}}^{\text{A}^3} &\approx \log_2 \frac{1}{b} + \int_0^\infty \log_2(1 + b\rho_s x) f_{\gamma_{\text{A}^3}}(x) dx \\
&\stackrel{(c_2)}{=} \log_2 \frac{1}{b} + \frac{1}{\ln 2} \sum_{i=1}^{NM} \sum_{j=1}^{NK} \mu_{i,NM} \mu_{j,NK} \left(e^{\frac{i\Omega_h + j\Omega_g}{b\rho}} \text{Ei} \left(-\frac{i\Omega_h + j\Omega_g}{b\rho} \right) - e^{\frac{i\Omega_h}{b\rho}} \text{Ei} \left(-\frac{i\Omega_h}{b\rho} \right) - e^{\frac{j\Omega_g}{b\rho}} \text{Ei} \left(-\frac{j\Omega_g}{b\rho} \right) \right) \\
&\stackrel{(c_3)}{\approx} \log_2 \frac{1}{b} + \frac{1}{\ln 2} \sum_{i=1}^{NM} \sum_{j=1}^{NK} \mu_{i,NM} \mu_{j,NK} \left(\ln \left(\frac{i\Omega_h + j\Omega_g}{b\rho} \right) - \ln \left(\frac{i\Omega_h}{b\rho} \right) - \ln \left(\frac{j\Omega_g}{b\rho} \right) - C \right) \\
&= \log_2 \frac{1}{b} + \frac{1}{\ln 2} \sum_{i=1}^{NM} \sum_{j=1}^{NK} \mu_{i,NM} \mu_{j,NK} \left(\ln \left(\frac{i\Omega_h + j\Omega_g}{ij\Omega_h\Omega_g} \right) + \ln(b\rho) - C \right) \\
&\stackrel{(c_4)}{=} \frac{1}{\ln 2} (\ln \rho - C) + \frac{1}{\ln 2} \sum_{i=1}^{NM} \sum_{j=1}^{NK} \mu_{i,NM} \mu_{j,NK} \ln \left(\frac{i\Omega_h + j\Omega_g}{ij\Omega_h\Omega_g} \right). \tag{48}
\end{aligned}$$

Specifically, the step (c₂) is obtained according to [25, Eq. (4.337.2)], the step (c₃) is obtained by applying the approximation $\text{Ei}(x) \stackrel{x \rightarrow 0}{\approx} C + \ln|x|$ [26] and the step (c₄) follows the property of the binomial coefficient $\sum_{i=0}^n (-1)^i \binom{n}{i} = 0$.

APPENDIX B

PROOF OF THE LEMMA 1

In the first stage of AIA-AS, the maximum elements h_i^{\max} and g_i^{\max} and their distributions are obtained according to (24)-(27). In the second stage of AIA-AS, the relatively smaller element $\gamma_i^w = \min(h_i^{\max}, g_i^{\max})$ in each row for $i \in \mathcal{N}$ is found out. Thus, the CDF of γ_i^w for $x \geq 0$ can be calculated as follows:

$$\begin{aligned}
F_{\gamma_i^w}(x) &= \Pr \{ \min(h_i^{\max}, g_i^{\max}) < x \} \\
&= \Pr (h_i^{\max} < g_i^{\max} < x) + \Pr (g_i^{\max} < h_i^{\max} < x) \\
&= \int_0^x f_{g_i^{\max}}(y) \int_0^y f_{h_i^{\max}}(z) dz dy + \int_0^x f_{h_i^{\max}}(y) \int_0^y f_{g_i^{\max}}(z) dz dy \\
&= 1 - \sum_{i=1}^M \sum_{j=1}^K \mu_{i,M} \mu_{j,K} e^{-(i\Omega_h + j\Omega_g)x}. \tag{49}
\end{aligned}$$

In the third stage of AIA-AS, $\gamma_{\text{AIA}}^w = \max(\gamma_i^w)$ for $i \in \mathcal{N}$ is selected. In the case that γ_{AIA}^w lies in the i_1^* row, we first define $\hat{\gamma}^w = \max_{i \neq i_1^*}(\gamma_i^w)$ and obtain the CDF of $\hat{\gamma}^w$ as follows:

$$F_{\hat{\gamma}^w}(x) = [F_{\gamma_{i_1^*}^w}(x)]^{N-1} \stackrel{(c_6)}{=} \sum_{\ell} C_{\ell} t_{\ell} e^{-\xi_{\ell} x}, \quad (50)$$

where the step (c_6) is expanded according to the Multinomial theorem. Specifically, $\ell_0 + \dots + \ell_{MK} = N - 1$, the multinomial coefficient $C_{\ell} = \binom{N-1}{\ell_0, \dots, \ell_{MK}} = \frac{(N-1)!}{\ell_0! \dots \ell_{MK}!}$, $t_{\ell} = \prod_{\substack{1 \leq i \leq M \\ 1 \leq j \leq K}} (-\mu_{i,M} \mu_{j,K})^{\ell_{ij}}$ and $\xi_{\ell} = \sum_{i=1}^M \sum_{j=1}^K (i\Omega_h + j\Omega_j) \ell_{ij}$.

Next we need to obtain the CDF and PDF of $\gamma_{\text{AIA}}^s = \max(h_{i_1^*}^{\max}, g_{i_1^*}^{\max})$ which lies in the same i_1^* row as γ_{AIA}^w . By applying some algebraic manipulations, we have

$$\begin{aligned} F_{\gamma_{\text{AIA}}^s}(x) &= \Pr \left\{ \max(h_{i_1^*}^{\max}, g_{i_1^*}^{\max}) < x, \gamma_{i_1^*}^w \geq \hat{\gamma}^w \right\} \\ &= N \left\{ \Pr(\hat{\gamma}^w < g_{i_1^*}^{\max} < h_{i_1^*}^{\max} < x) + \Pr(\hat{\gamma}^w < h_{i_1^*}^{\max} < g_{i_1^*}^{\max} < x) \right\}, \end{aligned} \quad (51)$$

and

$$f_{\gamma_{\text{AIA}}^s}(x) = N \left(f_{h_i^{\max}}(x) \int_0^x f_{g_i^{\max}}(y) \int_0^y f_{\hat{\gamma}^w}(z) dz dy + f_{g_i^{\max}}(x) \int_0^x f_{h_i^{\max}}(y) \int_0^y f_{\hat{\gamma}^w}(z) dz dy \right),$$

and by applying some algebraic operations, finally $f_{\gamma_{\text{AIA}}^s}(x)$ can be obtained as in (34).

APPENDIX C

PROOF OF THE PROPOSITION 2

Given the PDF of γ_{AIA}^s (i.e., $f_{\gamma_{\text{AIA}}^s}(x)$), we can approximate the average sum-rate of AIA-AS (i.e., $\bar{R}_{\text{sum}}^{\text{AIA}}$) according to (28) as below:

$$\begin{aligned} \bar{R}_{\text{sum}}^{\text{AIA}} &\approx \log_2 \frac{1}{b} + \int_0^{\infty} \log_2(1 + b\rho x) f_{\gamma_{\text{AIA}}^s}(x) dx \\ &= \log_2 \frac{1}{b} + \sum_{i=1}^M \sum_{j=1}^K \sum_{\ell} \frac{C_{\ell} t_{\ell}}{\ln 2} (T_1 + T_2 + T_3 + T_4), \end{aligned} \quad (52)$$

where T_1 is given by

$$\begin{aligned} T_1 &= - \int_0^{\infty} \frac{\xi_{\ell} \zeta_{ij} \ln(1 + b\rho x) e^{-j\Omega_g x}}{i\Omega_h \phi_i} dx \\ &\stackrel{(c_7)}{=} \frac{\xi_{\ell} \tilde{\zeta}_{ij}}{\phi_i} e^{\frac{j\Omega_g}{b\rho}} \text{Ei} \left(-\frac{j\Omega_g}{b\rho} \right) \stackrel{(c_8)}{\approx} \frac{\xi_{\ell} \tilde{\zeta}_{ij}}{\phi_i} \chi(j\Omega_g), \end{aligned} \quad (53)$$

in which $\tilde{\zeta}_{ij} = N\mu_{i,M}\mu_{j,K}$, $\phi_i = i\Omega_h + \xi_{\ell}$, $\chi(x) = C + \ln|\frac{x}{b\rho}|$, $\text{Ei}(x)$ is the Exponential integral function and $C \approx 0.577$ is the Euler's constant. Specifically, the step (c_7) is obtained with the help of [25, Eq.

(4.337.2)], and the step (c₈) is obtained by using the approximation of $e^x \approx 1$ and $\text{Ei}(x) \approx C + \ln x$ when $x \rightarrow 0$ [26]. Similarly,

$$\begin{aligned} T_2 &= - \int_0^\infty \frac{\xi_\ell \zeta_{ij} \ln(1 + b\rho x) e^{-i\Omega_h x}}{j\Omega_g \phi_j} dx = \frac{\xi_\ell \tilde{\zeta}_{ij}}{\phi_j} \chi(i\Omega_h), \\ T_3 &= - \int_0^\infty \frac{\zeta_{ij} \phi_{ij,2} \ln(1 + b\rho x) e^{-\phi_{ij,1} x}}{\phi_i \phi_j} dx = \frac{\zeta_{ij} \phi_{ij,2} \chi(\phi_{ij,1})}{\phi_i \phi_j \phi_{ij,1}}, \\ T_4 &= \int_0^\infty \frac{\zeta_{ij} (i\Omega_h + j\Omega_g) \ln(1 + b\rho x) e^{-(i\Omega_h + j\Omega_g)x}}{ij\Omega_h \Omega_g} dx = -\tilde{\zeta}_{ij} \chi(i\Omega_h + j\Omega_g), \end{aligned}$$

in which, $\phi_j = j\Omega_g + \xi_\ell$, $\phi_{ij,1} = i\Omega_h + j\Omega_g + \xi_\ell$, and $\phi_{ij,2} = i\Omega_h + j\Omega_g + 2\xi_\ell$.

APPENDIX D

PROOF OF THE PROPOSITION 3

Recall that the achievable rate of the secondary user in (11) consists of two terms. Here we can calculate the asymptotic average rate of UE₁ in PU-AS, which is denoted by $\bar{R}_1^P = \mathbb{E}[R_1^P]$, in two parts.

The first part of (11), which is denoted by $\bar{R}_{1,1}^P$, can be obtained by calculating the following integral for $\delta = 1$:

$$\begin{aligned} \bar{R}_{1,1}^P &= \int_0^\infty \int_0^x \log_2 \left(\frac{\rho x}{\varepsilon + 1} \right) f_{g^{\max}}(y) f_{h_i^{\max}}(x) dy dx \\ &= \sum_{i=1}^M \sum_{j=1}^{NK} \frac{i\Omega_h \mu_{i,M} \mu_{j,NK}}{\ln 2} \left\{ \int_0^\infty \ln \left(\frac{\rho x}{\varepsilon + 1} \right) e^{-i\Omega_h x} dx - \int_0^\infty \ln \left(\frac{\rho x}{\varepsilon + 1} \right) e^{-(i\Omega_h + j\Omega_g)x} dx \right\} \\ &\stackrel{y=\rho x/(\varepsilon+1)}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{i(\varepsilon + 1)\Omega_h \mu_{i,M} \mu_{j,NK}}{\rho \ln 2} \left\{ \int_0^\infty \ln y e^{-\frac{(\varepsilon+1)i\Omega_h y}{\rho}} dy - \int_0^\infty \ln y e^{-\frac{(\varepsilon+1)(i\Omega_h + j\Omega_g)y}{\rho}} dy \right\} \\ &\stackrel{(c_9)}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{\mu_{i,M} \mu_{j,NK}}{\ln 2} \left\{ -C - \ln \left(\frac{i(\varepsilon + 1)\Omega_h}{\rho} \right) + \frac{i\Omega_h C}{i\Omega_h + j\Omega_g} + \frac{i\Omega_h \ln \left(\frac{(\varepsilon+1)(i\Omega_h + j\Omega_g)}{\rho} \right)}{i\Omega_h + j\Omega_g} \right\} \\ &= \sum_{i=1}^M \sum_{j=1}^{NK} \frac{\mu_{i,M} \mu_{j,NK}}{\ln 2} \left\{ \frac{j\Omega_g \left(\ln \left(\frac{\rho}{\varepsilon+1} \right) - C \right) + i\Omega_h \ln(i\Omega_h + j\Omega_g)}{i\Omega_h + j\Omega_g} - \ln(i\Omega_h) \right\} \quad (54) \end{aligned}$$

Specifically, the step (c₉) is obtained with the help of [25, Eq. (4.331.1)].

The second part of (11), which is denoted by $\bar{R}_{1,2}^P$, can be obtained by calculating the following integral

for $\delta = 1$:

$$\begin{aligned}
\bar{R}_{1,2}^P &= \int_0^\infty \int_x^\infty \log_2 \left(\frac{\rho xy}{\varepsilon x + y} \right) f_{g^{\max}}(y) dy f_{h_i^{\max}}(x) dx \\
&= \underbrace{\int_0^\infty \int_x^\infty \log_2(\rho x) f_{g^{\max}}(y) dy f_{h_i^{\max}}(x) dx}_{T_5} + \underbrace{\int_0^\infty \int_x^\infty \log_2(y) f_{g^{\max}}(y) dy f_{h_i^{\max}}(x) dx}_{T_6} \\
&\quad - \underbrace{\int_0^\infty \int_x^\infty \log_2(\varepsilon x + y) f_{g^{\max}}(y) dy f_{h_i^{\max}}(x) dx}_{T_7}.
\end{aligned} \tag{55}$$

in which,

$$\begin{aligned}
T_5 &= \sum_{i=1}^M \sum_{j=1}^{NK} \frac{ij\Omega_h\Omega_g\mu_{i,M}\mu_{j,NK}}{\ln 2} \int_0^\infty \int_x^\infty \ln(\rho x) e^{-j\Omega_g y} dy e^{-i\Omega_h x} dx \\
&= \sum_{i=1}^M \sum_{j=1}^{NK} \frac{i\Omega_h\mu_{i,M}\mu_{j,NK}}{\ln 2} \int_0^\infty \ln(\rho x) e^{-(i\Omega_h + j\Omega_g)x} dx \\
&\stackrel{(c_{10})}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{\mu_{i,M}\mu_{j,NK}}{\ln 2} \left\{ \frac{-i\Omega_h (C + \ln(i\Omega_h + j\Omega_g) - \ln \rho)}{i\Omega_h + j\Omega_g} \right\},
\end{aligned} \tag{56}$$

where the step (c_{10}) is obtained with the help of the substitution $t = \rho x$ and that of [25, Eq. (4.331.1)].

Similarly,

$$\begin{aligned}
T_6 &= \sum_{i=1}^M \sum_{j=1}^{NK} \frac{ij\Omega_h\Omega_g\mu_{i,M}\mu_{j,NK}}{\ln 2} \int_0^\infty \int_x^\infty \ln y e^{-j\Omega_g y} dy e^{-i\Omega_h x} dx \\
&\stackrel{(c_9)}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{i\Omega_h\mu_{i,M}\mu_{j,NK}}{\ln 2} \int_0^\infty (\ln x e^{-j\Omega_g x} - \text{Ei}(-j\Omega_g x)) e^{-i\Omega_h x} dx \\
&\stackrel{(c_{10})}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{\mu_{i,M}\mu_{j,NK}}{\ln 2} \left(\ln(i\Omega_h + j\Omega_g) - \ln(j\Omega_g) - \frac{i\Omega_h (C + \ln(i\Omega_h + j\Omega_g))}{i\Omega_h + j\Omega_g} \right),
\end{aligned} \tag{57}$$

where the step (c_9) is calculated with the substitution $t = y/x$ and [25, Eq. (4.331.2)], and the step (c_{10}) is with [27, Eq. (2.5.3.1)]. For the T_7 , we have

$$\begin{aligned}
T_7 &= \sum_{i=1}^M \sum_{j=1}^{NK} \frac{ij\Omega_h\Omega_g\mu_{i,M}\mu_{j,NK}}{\ln 2} \int_0^\infty \int_x^\infty \ln(\varepsilon x + y) e^{-j\Omega_g y} dy e^{-i\Omega_h x} dx \\
&\stackrel{t=y-x}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{ij\Omega_h\Omega_g\mu_{i,M}\mu_{j,NK}}{\ln 2} \int_0^\infty \int_0^\infty \ln((\varepsilon + 1)x + t) e^{-j\Omega_g t} dt e^{-(i\Omega_h + j\Omega_g)x} dx \\
&\stackrel{(c_{11})}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{i\Omega_h\mu_{i,M}\mu_{j,NK}}{\ln 2} \int_0^\infty (\ln(\varepsilon + 1) + \ln x - \text{Ei}[-j\Omega_g(\varepsilon + 1)x] e^{j\Omega_g(\varepsilon + 1)x}) e^{-(i\Omega_h + j\Omega_g)x} dx \\
&\stackrel{(c_{12})}{=} \sum_{i=1}^M \sum_{j=1}^{NK} \frac{\mu_{i,M}\mu_{j,NK}}{\ln 2} \left(\frac{i\Omega_h (\ln(\varepsilon + 1) - C - \ln(i\Omega_h + j\Omega_g))}{i\Omega_h + j\Omega_g} + \frac{i\Omega_h (\ln(i\Omega_h + j\Omega_g) - \ln((\varepsilon + 1)j\Omega_g))}{i\Omega_h - \varepsilon j\Omega_g} \right)
\end{aligned} \tag{58}$$

where the step (c_{11}) is with [25, Eq. (4.337.1)] and the step (c_{12}) is with [25, Eq. (4.331.1)] and [27, Eq. (2.5.3.1)].

In this case, we can obtain $\bar{R}_{1,2}^P$ by summing up (56)-(58), and obtain \bar{R}_1^P by summing up (54) and (55). After some manipulations, \bar{R}_1^P can be obtained as in (46).

REFERENCES

- [1] Y. Yu, H. Chen, Y. Li, Z. Ding, and B. Vucetic, "Antenna selection for MIMO-NOMA networks", arXiv: 1609. 07978, 2016
- [2] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access", in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Jun. 2013.
- [3] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chin-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends", *IEEE Commun. Mag.*, vol. 53, pp. 74-81, Sep. 2015.
- [4] Z. Wei, J. Yuan, D. W. K. Ng, M. ElKashlan, and Z. Ding, "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks", arXiv: 1609. 01856, 2016.
- [5] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users", *IEEE Signal Process. Lett.*, vol. 21, pp. 1501-1505, Dec. 2014.
- [6] J. Choi, "Non-orthogonal multiple access in downlink coordinated two-point systems", *IEEE Commun. Lett.*, vol. 18, pp. 313-316, Jan. 2014.
- [7] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks", in *11th International Symposium on Wireless Commun. Systems (ISWCS)*, pp. 781-785, Aug. 2014.
- [8] J. Kim, and I. Lee, "Non-orthogonal multiple access in coordinated direct and relay transmission", *IEEE Commun. Lett.*, vol. 19, pp. 2037-2040, Aug. 2015.
- [9] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems", arXiv: 1607. 02668, 2016.
- [10] Y. Liu, Z. Ding, M. ElKashlan, and J. Yuan, "Non-orthogonal Multiple Access in Large-Scale Underlay Cognitive Radio Networks", *IEEE Trans. Veh. Technol.*, Mar. 2016
- [11] Z. Ding, F. Adachi, H. V. Poor, "The application of MIMO to non-orthogonal multiple access", *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, Jan. 2016.
- [12] Q. Sun, S. Han, I. Chin-Lin, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems", *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, Aug. 2015.
- [13] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback", arXiv: 1511. 05583, 2015.
- [14] B. Kim, S. Lim, H. Kim, S. Suh, J. Kwun, S. Choi, C. Lee, S. Lee, and D. Hong, "Non-orthogonal multiple access in a downlink multiuser beamforming system", in *IEEE Military Commun. Conf. (MILCOM)*, pp. 1278-1283, Nov. 2013.
- [15] Y. Liu, M. ElKashlan, Z. Ding, and G. K. Karagiannidis, "Fairness of user clustering in MIMO non-Orthogonal multiple access systems", *IEEE Commun. Lett.*, vol. 20, pp. 1465-1468, Apr. 2016.
- [16] A. F. Molish and M. Z. Win, "MIMO systems with antenna selection", *IEEE Micro. Mag.*, vol. 5, pp. 46-56, Mar. 2004.
- [17] R. Zhang and Y. C. Liang, "Exploiting Multi-Antennas for Opportunistic Spectrum Sharing in Cognitive Radio Networks", *IEEE J. Sel. Topics Signal Process.*, vol. 2, pp. 88-102, Feb. 2008.
- [18] S. Sanayei, and A. Nosratinia, "Antenna selction in MIMO systems", *IEEE Commun. Mag.*, vol. 42, pp. 68-73, Oct. 2004.
- [19] A. P. Shrestha, T. Han, Z. Bai, J. M. Kim, and K. S. Kwak, "Performance of transmit antenna selection in non-orthogonal multiple access for 5G systems", in *8th Int. Conf. on Ubiquitous and Future Netw. (ICUFN)*, Jul. 2016.

- [20] X. Liu and X. Wang, "Efficient antenna selection and user scheduling in 5G massive MIMO-NOMA system", in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, May 2016.
- [21] Z. Ding, P. Fan and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions", *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010-6023, Aug. 2016.
- [22] Z. Yang, Z. Ding, P. Fan, and Z. Ma, "Outage performance for dynamic power allocation in hybrid non-orthogonal multiple access systems", *IEEE Commun. Lett.*, vol. 20, pp. 1695-1698, Aug. 2016
- [23] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems", *IEEE Trans. Wireless Commun.*, Aug. 2016.
- [24] R. K. Jain, D. M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems", *DEC Technical Report 301*, Sept. 1984.
- [25] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 6th ed., Academic press, 2000.
- [26] W. J. Cody, and H. C. Thacher, "Chebyshev approximations for the exponential integral $Ei(x)$ ", *Mathematics of Computation*, pp. 289-303, 1969.
- [27] A. P. Prudnikov, Y. A. Brychkow, and O. I. Marichev, *Integrals and series (volume 2: special functions)*, Gordon and breach science publishers, 1983.